



Left Inverted Terminal Repeat: 1-103

Encapsidation Signal (Ψ): 183-331

HPRT Introns: 365-10547

Right Inverted Terminal Repeat:
10561-10663

pBR322 ori: 10867-11534

Kanamycin Resistance Gene: 12343-
13134

FIG. 1

SEQ ID NO: 1 pShuttle Sequence

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGG
GGGTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTGACG
TAGTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTGGCGGAACACATGTAA
GCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCCGGTGTACACAG
GAAGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTTGGGC
GTAAACCGAGTAAGATTTGGCCATTTTCGCGGGGAAAACCTGAATAAGAGGAA
GTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGTAATACTGGTACCGC
GGCCGCTCGAGTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTG
ATGGCTCTCAAATTCCTGCCTCCTTTAGGGATAAAAGACTTTAAGACTTT
TTAACAAAAAAGAAAAAGAAAAAATTCCTGCCTCCTGGTGTACACA
CACAGAAGGGTTCCCTCCCCTTGAATGTGACCAGGATCTGTGAAAATAAC
GGGATAGCCGCTCCTGTGATTAGGTTATGTGGTAGACTAGAGCAAGATTC
TCCTGCTGGTTTTGAAGAAGTCAGCTGCCATGTTGTGAGACTGTCATGGGC
TAGGGCATGAGCCTTTAAATATCTGGGAGCAACCCCTGGCCAGCAGCCAG
TGAGAAAACGGGCCCTCAGTCCTACAATCACAAGGAACCTAAATTCTGCCA
ACAACCTGAAGGAACCTTGAAGAGGATCATGAGTCCCTTGATTGAGCTTG
ATGAGCCCCTGAGCAGAGGATACAGCTAACTTGTACTAGGGAAGTATAAA
AAACATGCATGGGAATGATATATATCAACTTTAAGGATAATTGTCATACTT
CTGGGAATGAAGGGAAAGAAATGGGGCTTTAGTTGTATTATGATCTTTAA
TTTCTCAAAAAAATAAGATCAGAAGCAAATATGGCAAAATGTTAATACT
TTTGTGGGTACGTAGGTATTCAGCATACCCTTTTTTCTGAGTTCAAAATAT
TTTATAATTAAAATGAAATGCAGGCCAGGCACAGTGGCTCATGCCTATAA
TACCAGCACTTTGCGAGGCCGAGGTGGGAGGATGGCTTGAGGCCAGACCA
GCCTGGCCAACATGGCAAAACCCCATCTCTACTTAAAAAAAAAAAACTA
TATATATATATATGTGTGTGTGTGTATATATATATATGTATATATATTTA
TATATGTGTGTATATATATATATGTATATATATTTATATATGTGTGTGTATA
TATATATATACACACACACACATATATACATACATACATACACACACACA
CACACACAATTAGCCAGGCATGGTGGCGCACACCTGTAGTCCCAGCTACT
TGGGAGGCTGAGACATGAGAATTGCTTGAACCTGGGAGGCAGAGTAGTTA
GTGAGCTGAGATCATACCACTGCACTCCAGCCTGGTGACAGAGTGAGACT
CTGTCTTAAAAAAAAATAAAAATTTAAATTTAAATGCAAAAGGTCCAAGTGA
ATTGAAGAGGAAAGGGGTATCAAGGAAGGTTTTGTGGAGGTGACGTTTGA
GCTGGGTCTTAAATGACTTAAACATGGGATAAGAAGGGAGGGAATAAGG
ACATTTGAGGTACGAGAAATAAGGAGCAAACAGTGGAAACAACCTAACG
TCTGTCAACCAGTGAATGGATAACAAAAATGTAATTCAGATGGTATCCAA
CTTACGATGGTTCAACATGAGATTTTTCTGACTTTAGGATAGATTTATCAA
AGTAGTAAATCCATTTTCAACTTATGATATTTTCAACTTCAGATGGGTTTA
TCAGGACACAGTTGAGGAACACCTGTCTATCCATACAATTTGGCAATAAA
AAGGAAATGAGTGCAGATATACTCCACAACATGAATGAACCTTGAAAACA
TTAAGTGAGAGAAGCCAGATACAAAAGGCCACATATTGTATGATTCTATT
TATACAAAATGTCCAGAATAGGCAAATCTTATAGACAGCAAGTAGGTAGA
TGATCAGTTTGCTAGGTGCTGGGGGAAGGGGAAATGGGGAGTGATGGCTA
AGGGGATTGGGTTTCTTTGTGGGGCAATGAAAATGTTTTAAATTTAGCG
TGATAATGATTGCACAATGCTGCATATATATATAATCTATAGATTATATAT
ATATAAAGAGAGGCTGTTAGACAGTGATAAGTGATATATATATATATATA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGTTAGTGATAAGTGATC
AGGAAAATAAAAGTATTGAGGAGGAATACGAAGTTGACGGTGTGAAAAC
ATGAGATTTTATATAGGATGGCCAGGGAAGGCCTTAATGAGAAAGTGACT

FIG. 2

TATGAGTAAAAACAAGGGATCCTAAACCTTAGCATGCATCAGAATCACTC
GGAAACTTGTTAAAGCATAGCTTGCTGGGCCTCATCACAGATATTTTGATT
CGGTAGGTCTTGTCTGATATTAATACTTTTGGTCTAGGGAACCATTTT
GAGAACCCTGAGCTAAAGGAAGTAAAGGTTTCCCTTAGTTTACTAGCTG
GTAACCCTAGGAACTGCTTAGCCTCTCGGTGCTAAGATACAAAATACTTT
AGCACATAATAACACATGGAAAATAGTCTATAAAATTATAAATATTATTTT
TATGTACCAAATATTACATAAGACAAAATCTAAGCAAGATATATATATAT
ATACATAAAATATAAGATATATATGTATATATTATATATAGATAAATAGA
GAGAGAGAGTTATGTTTAGAAAGAAAATACTTCAAACCTAAAAAAGAGA
GGTAGGAAGTATACCATTCCATTATTGGTAAAAACAAATTACTAAGTAGT
CTTTACAAAAAACCAATCTCACTCCTTTAGAACACAAGCCCACCATTA
ACTGATGCAGAGGAATTTCTCTCCCTGGCTTACCTTTAGGATGGTGCATAC
TAAGTTAGAAAAGTCATAAATGTTATATTAAGTAAATGTGAACCTTACT
TCCACAATCAAGACATTCTAGAAGAAAAAGAGAAATGAAAATCAGTACA
ATGAATAAAACGGTATTTCCAATTATAAGTCAAATCACATCATAACAACC
CTAAGGAATTATCCAACTCTTGTTTTTAGATGCTTTATTATATCAAACCTCT
CCTTTAAACAAGTGGCCCATCTGCTGGGATTTGGAAGCCTGTAATACTGA
AATTTTCATCATAATGGAAATTTTAAAAACAGAATTTGACCCACCTGTTTT
TAAACACTTTTCATTACTTAACAAGAGGTCTAATCTTGGGCAAGTCTTGAA
ATTTCTCTGGCCTTAGTTTCCCATGTGTTAAATGAACTTGAAGCAGTTGG
TCTCTTATAGTCTCCTGACTCTAACATTCTAAGAATTATATTTGTACAATA
ACTCAAAAATCACATAATTTAATTTACCATATGGACTCCAAAATATATTTT
CTCATTAGGCTAAACTTGATCTGCATTTTCTGGATGTGTCCATATTCTTGG
ACTACACTAAACATGATACCAATGCTTCCTCTCACCATAAACCCCTCACTT
CGCTTTCTACATTTAAGAATTTTATAGCTGGAAGAGTCCTTAACAGAAAAT
ACCATCTAATAATTACCCCTCAAAATCGAGAAAGTCCTATCTGTTCTTATG
CTAGTTATAAGAATGAGGCAGCATTTTACATAATGGTTATAAACACTGCC
ACAAGAAGATTCATGATGTGTTGTTTATCTGTAGCTCTCATCATACTCTGT
CATATAACTATAGCATTAAAGATTTTAAATGTTCTATATATTCTTCTAAGACA
GTGTTTACCAGAGTAAGGCACAAAAGATCCACTGGTTTGCAAGAAAGATT
AGAACTTTTAAATTTTTTACCTCACCTTGTTTAAATCTATATTTTTGTATGTA
TTTTGTAACATATATATTATTATTACCATAAATCATATATAATTTAAATG
CATATATTAGGGGTAAATGCTCAGGAACTTTTTATAAATTGGGCATGCA
AATACAAGTTTGAAGACTCACTGTTCTAGGTATTAAAAGTAAAGTTATAA
CCAAGTAAAGCTTCCACCTTTTCATGTCTCAAAGCAGTTTATTGTTGGAGG
TAAGATCTCTTAGAAGCCTAAACAGGTCCAAGTACAGAAATGAAGTAAGGC
TAGCCCATAACTTGTGGCAAGCAATTCATACTATTTCTCTCATGCTGAGCT
CTCCTCAGTGAAGCAGCTACTATAGACAACTGCAGCCTATTGGTAGCCTAT
TTTACAGGCAGGAAAAAAATTACTTTTTATTCAAAGTGGAACCTCAGGACA
TGGGGAGAAAAATGAATACAAAAAATAGGGTCAATCCAAAGGCACACAGC
AAATGAGTAACACAGTTATGTTTTTTTCCCATTTGTATGAGGTCCCAGTAA
ATTCTAAGTAACTGCAAATTTAATAATACACTAAAAAAGCCATGCAATT
GTTCAAATGAATCCCAGCATGGTACAAGGAGTACAGACACTAGAGTCTAA
AAAACAAAAGAATGCCATTATTGAGTTTTTGAATTATATCAAGTAGTTAC
ATCTCTACTTAATAAATGAGAAAAACGAGGATAAGAGGCCATTTGATAAA
ATGAAAATAGCCAAGAAGTGGTATTAGAGACTTGAATACAGGTATTCGGG
TCCAAAGTTCATCTGCTCAAATACTAACTGGGGAAAAGAGGGAAAAATAT
TTATATACATATATATCTGCACACAAAAATACCCCCAAAAGACAAAATGA
GGCCAGGCAGGGTGGCTCACACCCGTAATCCCGGTACTTTGGGAGGCTGA
GGCAGGTGGATACCTGAGATCAGGAGTTGGAGATCAGCCTGGTCAACATG

FIG. 2a

GTGAAACCCTGTCTCTACTAAAGATAAAAAAATTAGCCAGGCATGGTGGC
 GTGCGCCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGAGAATCACTT
 GAACTGGGAAGGGGAGGTTGCAGTGAGCCAAGATCGTACTACTGCACTCC
 AGCCTGGGCAGCAGAGTGAGACTCCATCACAAAAATAAATAAATAAATA
 AAATACAATGAAACAGAAAGTTCAAATAATCCCATAATCTTACCACCAAG
 AAATAACTTTCACTCGTTATACTTATTGATTTTTTCCATAATAAATGTACTTT
 ACTGTGACTATCATGAAAAGAAAGTTATTTTAGAAACAGAGAACTGTTTC
 AGATCAAATCTATGTAGTAGAACAGAGCCATTAGGTGGGAAAGACGAGA
 TCAAACCTAAATCTCAGAAGGCCTAAAAGGCTAGGTCCATTCCAGCACTAA
 AAACCTGACCAGACAAGTAATGGCTTCAACAGCTTCTAAATATGGACAAAG
 CATGCTGAAAGGGAAGGACAGGTCTAACAGTGGTATATGAAATGAACAG
 GAGGGGCAAAGCTCATTTCTCCTCTGAAGTTTTCCAAAGATGCTGAGGAG
 GACATTAGTTTGACATGACCCTGATATGGGACAAGATAATTTACAGAAAG
 TTTTACATGTTAAAGTTTTCTTATAGATACTCATTCAAGTAAGCAATGAAC
 ACTAAAATCTAAAGAAAGAAAAGAGCTTTAGAGTCAGGTCTGTATTCAA
 TTCAAGCTCTACCACTTACTGGTTCTGTGACTTTGGGCAAGTCTTTTAACCT
 TATTAAGTCTTAATTTCTGATTTGTAAAATGGGGATATCGTCTCCCTCAC
 AGGATTGTTGTGAAACTTTTATGAGATTAATGCCTTTATATTTGGCATAGT
 GTAAGTAAACAATAACTGGCAGCTTCAAAAAAAAAAAGCAGTAGCATTCC
 ATCATTTATTATTGGTTACTCTCAAAAAGTTTTTCAATGTACTAGAAGATA
 AATATTCAAATACCTTAATATCTCCATTATTTTCAGGTAAACAGCATGCTC
 CTGAACAACCAATGGGTCAACAAATAAATTAAAAGGGAAATCTAAAAAC
 ATCTTGATATTAACTACATGGAAGCACAATATACCAAACCAATGGTTC
 AACTAGGAGAATTTTAAGGTACAAGAAAACCTTTGAGATTTCTTAAAA
 TAATAGTATGTCTGAATTTATTGAGTGATTACCAGAACTGTTGTAAGAG
 CTCTACTTGCATTATAGCACTTAATCCTCTTAACCTCTATGGCTGCTATTATC
 AACCTCACCTAATCACATATGGGACACAGAGAGGTAAAGTAACTTGCCC
 AAGGTCAGAGTTAGGAAGTACTAAGCCATGCTTTGAATCAGTTGTCAGGC
 TCCGGAACCTCACACTTTCAGCCACTACATAATACTGCTTTGCTATCTTTTA
 GGAAACTATGTGAGTCTACCTCACATAGACTCACATAGGTTTGTTTTTTTT
 TTTTTTTTAAAGGCTATCTTTTCCCCCATCAATGTTTTTTGAAGGATCCCAA
 ATTAGAGTCCCACAGAGGCAGACAGCAGTACTTGACAATATGGACATTTA
 AGGTAAATGTTGGATTCTACTGTCTTTTTACTACATGACCTAGGGAACGAT
 AATTAACCTAGACTGCTTCCAAGGGTTAAATAACCCATTAGTTATACTAT
 GTAAATTATCTCTTAGTGATTGATTGAAAGCACACTGTTACTAATTGACTC
 GGTATGAAGTGCTTTTTTTCTTCCCTTTCAAGATACATACCTTTCCAGTTA
 AAGTTGAGAGATCATCTCCACCAATTACTTTTATGTCCCCTGTTGACTGGT
 CATTCTAGTTAAAAAAAAAAAAAAAACTATATATATATATATCTACACACAC
 ATATGTATATGTATATCCTTATGTACACACACAACTTCAAATTAATGAG
 AACTAGAAGATTTGAGAAGTTAGCTAGCTAATATCCATAGCATTATGATA
 TTCTAAATGATATGAATTATAAGAATTAGGTTTCCTGAAATGAATGACTAG
 AAAACTTTCAAGTAGAGATTAGTAAAAATTAAGAGTCTAATCGGCCAT
 TACTGATTTGATGTTTTTAAGAGTCCTAAAAAATGGGTTACATCCATTTTT
 AAGTGGGTTAGTATTATAACAGCCACCCATCTTCAATCACAGTGATTTCTGA
 ATTGTGAGGGAAGTTATTAGCATGACAGGTGTCTGGTTCTGGCCCTGTACG
 ATTCCCATGAGTCAAGCAAATTGTAAGGGCTGGTCTATATCACACCCAAC
 CCAAGGATATGTCCCTCAAAGTCTAGCCCAGGCCCGTCATCTTCAGC
 ATCATCTGGGAAACCAGGTCTGATTAGTAGTCCTTTAAGGAATACCTCTTA
 GGCTCCCATTTTACTGCTATCACAGAATCCAATAAAACCCTTACAGGAGAT
 TCAATGGGAAATGCTCAACACCCACTGTAGTTGGTGGTGACAATGACCAT

FIG. 2b

AATTTGGCTGTGCTGGATTCAGGACAGAAAATTTGGGTGAAAGAGCAGGT
GAACAAAAGAGCTTCGACTTGCCCTAGCAGAGAGCAAGCCATACCATAACC
ACAAAGCCACAGCAATTACAACGGTGCAGTACCAGCACAGTAAATGAAC
AAAGTAGAGCCCAGAAACAGACCCAGAACTATATGAGGATTTAGTATACA
ATAAAGATGGTATTTTCGAGTCAGTAGGGAAAAGATGAATTATTCAATAAA
TGATGTTTGGCCAACTAGTAACCCATTTGGGAAAAAATAAAAAGTATGGTC
CCTACCTCACAGCATACACAAAAATAAATTCCAGACGGATTAAAAATCTAA
ATGTAAAAAATAAAGCCATAAGTGGACTGGAAGAAAATAGAGAATTTTTT
TTAACATCCGTAGAAAGGGTAAAAACCCAGGCATGACATGAACCAAACT
GAAGAGGTTCTGTAACAAATACCCCCTTTTATATATTGGGCTCCAACAATA
AGAACCCATAGGAAAATGGAGAATGAACACAAATAGACAATTTATAGAA
GAGAAGGTTATAAGGTGTAAAATTATATCTATCTGAGAAACAAACACTAA
AACAAATGTGATTCTACTGTTCTCCACCCATACTGGCAAACTTAAGCCTG
ATAATATGCTGAGGGGAAATAAGCACTCTTGTTGGTGAGAGTATTAATTG
GCATAGCTTCTTTTGAAAATGACATAGCAATACCTGTTAAAATTGCAAAC
ATGCATGTCACTTAATCCAGTAATCCCCTTCTGGGAATCAATGCTACAAA
AACACTGACAAGTATACAAAGATACATTCAAGAGTGTTCACTGGGCCGGG
TGCGGTGGCTTCATGCCTGTAATCCCAGGGAGGCAGAGGCAAGACGATCG
CTTGACCCCAAGGAGTTCAAGGCCAGCCCGAGAAACACAGCAAGACCCTGT
CTCTCTTTTTTTTATTTAAAAAATAAATGTTCACTGTATCAGTTGTTCACAA
AAACAAACCAACATGTCCATTAACAGGGAACCATTTAAATTAATCAAGTT
CATCTACACAATGTAATACCATGCAACTATTTAAAAAGCACCTGATAATCC
AAAGCACACTGAGACAGAATAATGCTATTTAAAAACACCAAGTAGTGGA
CACTGTGTTGCCTATGACACCATTTTTATTCAACATTTAAACAAATTTGTA
ACAGCAATTACATGAGTAGTGACAATGGCGTTTATGAGACTTTTCACTTTT
ATGTGCTTCTATTTTTGTTATGCTTCTATATATACATCCATTTATTATGGAG
TGTTACTTTCAAAAATCACAAATGGGCCAGTATTATTTGGTGTTGCAAGGT
GAGCATATGACTTCTGATATCAACCTTTGCATATTACTTCTCAATTTAGGG
AAATTACAGACATCCCTTATTCTAACTAACTTAAAACCCAGCATTTCAAAC
ATACAGAATTGATGGGGAAAAAAAAGAAAGAAGAAAGAAAGAAAGGC
AACAAAGCTTCAGATGACAGTGACTCACATCAAATTATTTATAAAATCTGTT
AAATAGTGCCATCTTCTGGAGATACCTGGTATTACAGTCCAACCTCCAGTTG
ATGTCTTTACAGAGACAAGAGGAATAAAGGAAAAAATATTCAAGAACTG
AAAAGTATGGAGTCATGGAAAAATTGCTGTGATCCAAAGGCTACGGTGAT
AGGACAAGAAACAAGAGAACTCCAAGCAGTAAGACACTGCTGTTCTATTA
GCATCCAAACCTCCATACTCCTGTTTGCCCCAAGGCTTTTTTAAAAAATAG
AGACAGGATCTCACTATTTTGCTCAGGCTGGTCTTGAACCTCCTGGACTCAA
GCTATCCTCCTGCCTCGGCCTCCTAAAGTGCCGAGATTACAGGCTTGAGTC
ACCATACCTGGCTATTTATTTTTTCTTAACTCTCTTGCTGGCCTATAGCCA
CCATGGAAGCTAATAAAGAATATTAATTTAAGAGTAATGGTATAGTTCAC
TACATTGGAATACAGGTATAAGTGCCTACATTGTACATGAATGGCATAACA
TGGATCAATTACCCACCTGGGTGGCCAAAGGAAGTGCAGCAACCTCCCT
CCTTGGCTGTCTGGAACAAGCTTCCCACTAGATCCCTTTACTGAGTGCCTC
CCTCATCTTTAATTATGGTTAAGTCTAGGATAACAGGACTGGCAAAGGTG
AGGGGAAAGCTTCCTCCAGAGTTGCTCTACCTCTCCTCTACCGTCCTATC
TCCTCACTCCTCTCAGCCAAGGAGTCCAATCTGTCCTGAACCTCAGAGCGTC
ACTGTCAACTACATAAAATTGCCAGAGAAGCTCTTTGGGACTACAAACAC
ATACCTTAATGTCTTTATTTCTATTTTGTCTACCTCTTCAGTCTAGGTGAA
AAAATAGGAAGGATAATAGGGAAGAACTTTGTTTATGCCTACTTATCCGC
CCCTAGGAATTTTGAAAACCTCTAGGTAGCAATAAGAACTGCAGCATGGT

FIG. 2c

ATAGAAAAAGAGGAGGAAAGCTGTATAGAAATGCATAATAAATGGGCAG
 GAAAAGAAGCTTGCTTGGAACAAACAGGGAGGTTGAACTATAAGGAGAGAA
 AGCAGAGAGGCTAATCAACAAGGCTGGGTTCCCAAGAGGGCATGATGAG
 ACTATTACTAAGGTAGGAATTACTAAGGGCTCCATGTCCCCTTAGTGGCTT
 AGTACTATGTAGCTTGCTTTCTGCAGTGAACCTCAGACCCTTCTTTTAGGA
 TCCTAGAATGGACTTTTTTTTTTTATCGGAAAACAGTCATTCTCTCAACATT
 CAAGCAGGCCCAAGTCTACCACACTCAATCACATTTTCTCTTCATATCAT
 AATCTCTCAACCATTCTCTGTCTTTTAACTGTTTTTCTATACCCTGATCAA
 ATGCCAACAAAAGTGAGAATGTTAGAATCATGTATTTTATAGAGGTAGACT
 GTATCTCAGATAAAAAAAAAAAGGGCAGATATTCCATTTTCCAAAATATGTA
 TGCAGAAAAAATAAGTATGAAAGGACATATGCTCAGGTAACAAGTTAATT
 TGTTTACTTGTATTTTATGAATTCCCTAAAACCTACGTCACCCGCCCGTTC
 CCACGCCCGCGCCACGTCACAACTCCACCCCTCATTATCATATTGGCT
 TCAATCCAAAATAAGGTATATTATTGATGATGTTAATTAACATGCATGGAT
 CCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCA
 TCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTT
 GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
 CTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
 CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
 CCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGT
 ATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
 CCCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAAGTATCGTCTTGAGT
 CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
 GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT
 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
 AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTA
 CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG
 TCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGAG
 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTT
 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
 GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTATCCA
 TAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTA
 CCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC
 TCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA
 AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGG
 GAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
 CATTGCTGCAGCCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT
 CACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCA
 GCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGCA
 GGTAGCTTGCAAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTAT
 GGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTT
 GGGAAGCCCTGCAAAGTAACTGGATGGCTTTCTTGCCGCCAAGGATCTG
 ATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTC
 GCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTG
 GAGAGGCTATTCCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA
 TGCCGCCGTGTTCCGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTGTCAA

FIG. 2d

GACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGC
TATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTG
TCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCA
GGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGC
TGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGA
CCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCC
GGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCC
AGCCGAACTGTTTCGCCAGGCTCAAGGCGAGCATGCCCCGACGGCGAGGATC
TCGTGCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAATA
GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC
TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGG
CGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTC
GCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTTTGTTA
AAATTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAA
AATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTT
CAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAA
GGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC
CTAATCAAGTTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACC
CTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGT
GGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCT
GGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTA
ATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGATCGAATTAATTCT
TAATTAA

FIG. 2e

SEQ ID NO:2 Human TM amino acid sequence

MLGVLVLGALALAGLGFPAPAEPPGGSQCV EHD CFALYPGPAT
FLNASQICDGLRGHLM TVRSSVAADVISLLLNGDGGVGRRRLWIGLQLPPGCGDPKR
LGPLRGFQWVTGDNNTSYSRWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQQ
CEVKADGFLCEFHF PATCRPLAVEPGAAAAAVSITYGTPFAARGAD FQALPVGSSAA
VAPLGLQLMCTAPPGAVQGHWAREAPGA WD C SVENGGCEHACNAIPGAPRCQCPA
GAALQADGRSCTASATQSCNDLCEHFCVPNPDQPGSYSCMCETGYRLAADQHRCE
VDDCILEPSPCPQRCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQCQPLN
QTSYLCVCAEGFAPIPHEPHRCQMFCNQTACPADCDPNTQASCECPEGYILDDGFICT
DIDECENGGFCSGVCHNLPGT FECICGPDSALARHIGTDCDSGKVDGGDSGSGEPPPS
PTPGSTLTPPAVGLVHSGLLIGISIASLCLV VALLALLCHLRKKQGAARAKMEYKCAA
PSKEVVLQHV RTERTPQRL

FIG. 3

SEQ ID NO:3 human TM nucleotide sequence

atgcttggg gtcctgggcc ttggcgcgct ggccctggcc ggccctgggt tccccgcacc cgcagagccg cagccgggtg
gcagccaagt cgtcgagcac gactgcttcg cgctctaccc gggccccgcg accctcctca atgccagta gatctgcgac
ggactgcggg gccacctaata gacagtgcgc tcctcggtgg ctgccgatgt catttccttg ctactaacg gcgacggcgg
cgttggccgc cggcgccctt ggatcggcct gcagctgcca cccggctgcg gcgaccccaa gcgctcggg cccctgcgcg
gcttccagt gggtacggga gacaacaaca ccagctatag caggtgggca cggctcgacc tcaatggggc tccccctgc
ggccccgtgt gcgtcgctgt ctccgctgt gaggccactg tgcccagcga gccgatctgg gaggagcagc agtgcgaagt
gaaggccgat ggcttctct gcgagttcca ctccagcc accctgcaggc cactggctgt ggagcccgc gccgcggctg
ccgcccgtc gatcacctac ggcaccccgt tcgcggcccg cggagcggac ttccaggcgc tgccggtggg cagctccgcc
gcggtggctc cctcggctt acagctaag tgcaccgcgc cggccggagc ggtccagggg cactgggcca gggaggcgcc
gggcgcttg gactgcagc tggagaacgg cggctgcgag cacgcgtgca atgcgatccc tggggctccc cgctgccagt
gcccagccgg cggcgccctg caggcagacg ggcgctcctg caccgcatcc gcgacgcagt cctgcaacga cctctgcgag
cacttctgcg tccccaaacc cgaccagccg ggctcctact cgtgcatgtg cgagaccggc taccggtgg cggccgacca
acaccggtgc gaggacgtgg atgactgcat actggagccc agtccgtgtc cgcagcgtg tgtcaacaca cagggtggct
tcgagtcca ctgtaccct aactacgacc tggtagcagg cgagtgtgtg gagccgtgg acccgtgctt cagagccaac
tgcgagtacc agtgcagcc cctgaacaa actagctacc tctgcgtctg cggcgaggc ttgcgcca tccccacga
gccgcacagg tgccagatgt tttgaacca gactgcctgt ccagccgact gcgacccaa caccaggct agctgtgagt
gccctgaagg ctacatctg gacgacggtt tcatctgcac ggacatcgac gagtgcgaaa acggcggtt ctgtccggg
gtgtgccaca acctccccg tacctcgag tgcattgcg ggcccgactc ggcccttgc cggcacattg gcaccgactg
tgactccggc aagggtggac gtggcgacag cggctctggc gagccccgc ccagcccgc gcccggtcc acctgactc
ctccggccgt ggggctcgt cattcgggt tgcctatagg catctccatc gcgagcctgt gcctggtgtt ggcgctttg
gcgctcctt gccacctgc caagaagcag ggcgccgcca gggccaagat ggagtacaag tgcgcggccc ctccaagga
ggtagtgct cagcacgtgc ggaccgagc gacgccgcag agactc

FIG. 4

SEQ ID NO: 4

GTTTAAACGGGCCCCTCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTAC
GGGGTCATTAGTTCATAGCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGG
CCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTTGCC
AGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG
CCTGGCATTATGCCAGTNCATGACCTTATGGGACTTTCTACTTGGCAGACATCTACGTATTAGTC
ATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCA
CGGGGATTTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACG
GGACTTTTCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGT
GGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCCTGCTTACTGGCTTATCGAGAT
ATCTGCAGAATTTCATCTGTCTGACTGCTACCGGCAGCGCGCAAGAAGTGTCTGGGCTGGG
ACGGACAGGAGAGGCTGTGCGCATCGCGTCTGTGCCCTCTGCTCCGGCACGCAATCTGTCGAG
TGCCCGCGCTTTCCCGGCGCCTGCACGCGGCGCGCTGGGTAACATGCTTGGGGTCTGGTCTCT
GGCGCGCTGGCCCTGGCCGGCCTGGGGTTCCCCGCACCCGCAGAGCCGCAGCCGGGTGGCAGCCA
GTGCGTCGAGCACGACTGCTTCGCGCTCTACCCGGGCCCCGCGACCTTCTCAATGCCAGTCAGAT
CTGCGACGGACTGCGGGGCCACCTAATGACAGTGCCTCTCGGTGGCTGCCGATGTCATTTCTCT
GCTACTGAACGGCGACGGCGGCGTTGGCCGCCGGCGCCTCTGGATCGGCCTGCAGCTGCCACCCG
GCTGCGGCGACCCCAAGCGCCTCGGGCCCCCTGCGCGGCTTCCAGTGGGTTACGGGAGACAACAAC
ACCAGCTATAGCAGGTGGGCACGGCTCGACCTCAATGGGGCTCCCCTCTGCGGCCCGTTGTGCGTC
GCTGTCTCCGCTGCTGAGGCCACTGTGCCACGAGCCGATCTGGGAGGAGCAGCAGTGCAGAACT
GAAGGCCGATGGCTTCTCTGCGAGTTCCACTTCCAGCCACTGCAGGCCACTGGCTGTGGAGCC
CGGCGCCGCGGCTGCCGCGTCTCGATCACCTACGGCACCCCGTTGCGGGCCCGGAGCGGACTT
CCAGGCGCTGCCGCTGGGCAGCTCCGCGCGGCTGGCTCCCCTCGGCTTACAGCTAATGTGCACCGC
GCCGCCCGGAGCGGTCCAGGGCACTGGGCCAGGGAGGCGCCGGGCGCTTGGGACTGCAGCGTG
GAGAACGGCGGCTGCGAGCACGCGTGCAATGCGATCCCTGGGGTCCCCGCTGCCAGTGCCACAGC
CGGCGCCGCCCTGCAGGCAGACGGGCGCTCTGCACCGCATCCGCGACGCACTCTGCAACGACC
TCTGCGAGCACTTCTGCGTTCCCAACCCCGACCGCCGGGCTCTACTCGTGCATGTGCGAGACCG
GCTACCGGCTGGCGGCCGACCAACACCGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCCAGT
CCGTGTCCGCAGCGCTGTGTCAACACACAGGGTGGCTTCGAGTGCCACTGCTACCCTAACTACGAC
CTGGTGGACGGCGAGTGTGTGGAGCCCCGTGGACCCGTGCTTCAGAGCCAACTGCGAGTACCAGTG
CCAGCCCCCTGAACCAAACTAGCTACCTCTGCGTCTGCGCCGAGGGCTTCGCGCCCATTTCCCAACGA
GCCGACAGGTGCCAGATGTTTTGCAACCAAGACTGCGCTGTCCAGCCGACTGCGACCCCAACCCCA
GGCTAGCTGTGAGTGCCCTGAAGGCTACATCCTGGACGACGGTTTCATCTGCACGGACATCGACGA
GTGCGAAAACGGCGGCTTCTGCTCCGGGGTGTGCCACAACTCCCCGGTACCTTCGAGTGCATCTG
CGGGCCCGACTCGGCCCTTGCCCCGCCACATTGGCACCGACTGTGACTCCGGCAAGGTGGACGGTG
GCGACAGCGGCTCTGGCGAGCCCCCGCCAGCCGACGCCCCGGTCCACCTTGACTCCTCCGGCCG
TGGGGCTCGTGCAATTCGGGCTTGCTCATAGGCATCTCCATCGCGAGCCTGTGCCTGGTGGTGGCGC
TTTTGGCGCTCCTCTGCCACCTGCGCAAGAAGCAGGGCGCCGCCAGGGCCAAGATGGAGTACAAG
TGCGCGGCCCTTCCAAGGAGGTAGTGCTGCAGCACGTGCGGACCGAGCGGACGCGCAGAGACT
CTGAGCGGCCTCCGTCCAGGAGCCTGGCTCCGTCCAGGAGCCTGTGCCTCCTACCCCCAGCTTTG
CTACCAAAAGCACCTTAGCTGGCATTACAGCTGGAGAAGACCCCTCCCCGCACCCCCCAAGCTGTTTT
CTTCTATTCCATGGCTAACTGGCGAGGGGTGATTAGAGGGAGGAGAATGAGCCTCGGCCTCTTCC
GTGACGTCACTGGACCACTGGGCAATGATGGCAATTTTGTAAACGAAGACACAGACTGCGATTTGTC
CCAGGTCCTCACTACCGGGCGCAGGAGGGTGAGCGTTATTGGTCGGCAGCCTTCTGGGCAGACCTT
GACCTCGTGGGCTAGGGATGACTAAAATATTTATTTTTTTTAAAGTATTTAGGTTTTTGTGTTTCTCT
TTGTTCTTACCTGTATGTCTCCAGTATCCACTTTGCACAGCTCTCCGGTCTCTCTCTCTACAACT
CCCACTTGTCATGTGACAGGTAACTATCTTGGTGAATTTTTTTTCTAGCCCTCTCACATTTATG
AAGCAAGCCCCACTTATCCCCATTCTTCTAGTTTTCTCTCTCCAGGAACTGGGCCAACTCACCTG
AGTACCCCTACCTGTGCCTGACCCTACTTCTTTTGTCTTAGCTGTCTGCTCAGACAGAACCCCTAC
ATGAAAACAGAAACAAAAACACTAAAAATAAAAAATGGCCATTTGCTTTTTTACCAGATTTGCTAATT
TATCCTGAAATTTAGATTCCCAGAGCAAAATAATTTTAAACAAAGGTTGAGATGTAAAAGGTATT
AAATTGATGTTGCTGGACTGTATGAAATTAACCCAAAGAGGTATTTATCTTTACTTTTAAACA
GTGAGCCTGAATTTTGTGCTGTTTTGATTGTACTGAAAAATGGTAATTGTTGTAATCTTCTTAT
GCAATTTCTTTTTTGTATTATTACTTATTTTTGACAGTGTTGAAAATGTTTCAAGAGGTTGCTCTAG
ATTGAGAGAAGAGACAAACACCTCCCAGGAGACAGTTCAAGAAAGCTTCAAACCTGCATGATTCAT
GCCAATTAGCAATTGACTGTCACTGTTCTTGTCACTGGTAGACCAAAATAAAACCAGCTCTACTG
GTCTTGTGGAATTGGGAGCTTGGGAATGGATCCTGGAGGATGCCCAATTAGGGCCTAGCCTTAATC

FIG. 5

AGGTCCTCAGAGAATTTCTACCATTTTCAGAGAGGCCTTTTGGAATGTGGCCCCTGAACAAGAATTG
GAAGCTGCCCTGCCCATGGGAGCTGGTTAGAAATGCAGAATCCTAGGCTCCACCCCATCCAGTTCA
TGAGAATCTATATTTAACAAGATCTGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATT
CAGACTGCTTCCAATTTTCTGGAATACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAG
GCCCTTATTTTCAAGAACTGAGGAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGG
TACACAGCTCTAGACACTGCCACACAGGGTCTGCAAGGTCTTTGGTTCAGCTAAGCTAGGAATGAA
ATCCTGCTTCAGTGTATGGAAATAAATGTATCATAGAAATGTAACCTTTTGTAAAGACAAAGGTTTT
CTCTTCTATTTTGTAAACTCAAAATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACA
CAGGCAAAATCCTTGCTTATGACATCACTTGTACAAAATAAACAAATAACAATGTGAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGTAGCAGTCGACAGATGAATTCCACCACACTG
GACTAGTGGATCCGAGCTCGGTACCAAGCTTAAGTTTAAAC

FIG. 5a

SEQ ID NO 5

TCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTA
TCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTNCATGACCTTATGGGACTTTCCTACTTGGCAGACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAG
TCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCCTGCTTACTGGCTTATCGAGATATC

FIG. 6

SEQ ID NO 6

GGCAGCGCGCAGCGGCAAGAAGTGTCTGGGCTGGGACGGACAGGAGAGGCTGTGCCATCGGCG
TCCTGTGCCCCCTCTGCTCCGGCACGGCCCTGTGCGAGTGCCCGCGCTTTCCCCGGCGCCTGCACGC
GGCGCGCCTGGGTAACATGCTTGGGGTCTGTGCTTGGCGCGCTGGCCCTGGCCGGCCTGGGGTT
CCCCGCACCCGCAGAGCCGCAGCCGGGTGGCAGCCAGTGCCTCGAGCACGACTGCTTCGCGCTCT
ACCCGGGCCCCGCGACCTTCCTCAATGCCAGTCAGATCTGCGACGGAAGTGGGGGCCACCTAATG
ACAGTGCCTCTCGGTGGCTGCCGATGTCAATTTCTTGTCTACTGAACGGCGACGGCGGCGTTGGC
CGCCGGCGCCTCTGGATCGGCCTGCAGCTGCCACCCGGCTGCGGCGACCCCAAGCGCCTCGGGCC
CCTGCGCGGCTTCCAGTGGGTTACGGGAGACAACAACACCAGCTATAGCAGGTGGGCACGGCTCG
ACCTCAATGGGGCTCCCTCTGCGGCCCGTTGTGCGTCGCTGTCTCCGCTGCTGAGGCCACTGTGC
CCAGCGAGCCGATCTGGGAGGAGCAGCAGTGCAGAAAGTGAAGGCCGATGGCTTCCTCTGCGAGTTC
CACTTCCCAGCCACCTGCAGGCCACTGGCTGTGGAGCCCCGGCGCCGGCTGCCGCCGTCTCGATC
ACCTACGGCACCCCGTTTCGCGGCCCGCGGAGCGGACTTCCAGGCGCTGCCGGTGGGCAGCTCCGC
CGCGGTGGCTCCCCCTCGGCTTACAGTCAATGTGACCGCGCCCGCGGAGCGGCTCCAGGGGCACT
GGGCGAGGAGGGCGCGCGGCTTGGGACTGCGAGCTGGAGAACGGCGGCTGCGAGCAGCGCTG
CAATGCGATCCCTGGGGCTCCCGCTGCCAGTGCCAGCCGGCGCCCGCTGCAGGCAGACGGGC
GCTCCTGCACCGCATCCGCGACGCAGTCTGCAACGACCTCTGCGAGCACTTCTGCGTTCCCAACC
CCGACCAGCCGGGCTCCTACTCGTGCATGTGCGAGACCGGCTACCGGCTGGCGGCCGACCAACAC
CGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCAGTCCGTGTCCGCAGCGCTGTGTCAACAC
ACAGGGTGGCTTCGAGTGCCACTGCTACCCTAAGTACGACCTGGTGGACGGCGAGTGTGTGGAGC
CCGTGGACCCGTGCTTCAGAGCCAAGTGCAGTACCAGTGCAGCCCTGAACCAAAGTACGTAC
CTCTGCGTCTGCGCCGAGGGCTTCGCGCCCATTCACACGAGCCGCACAGGTGCCAGATGTTTTGC
AACCAGACTGCCTGTCCAGCCGACTGCGACCCCAACACCCAGGCTAGCTGTGAGTGCCCTGAAGG
CTACATCTGGACGACGGTTTCATCTGACAGGACATCGACGAGTGCGAAAACGGCGGCTTCTGCTC
CGGGTGTGCCACAACCTCCCGGTACCTTCGAGTGCATCTGCGGGCCCGACTCGGCCCTTGCCCG
CCACATTGGCACCGACTGTGACTCCGGCAAGGTGGACGGTGGCGACAGCGGCTCTGGCGAGCCCC
CGCCAGCCCGACGCCCCGGCTCCACCTTGACTCCTCCGGCCGTGGGGCTCGTGCATTGCGGCTTGC
TCATAGGCATCTCCATCGCGAGCCTGTGCTGGTGGTGGCGCTTTTGGCGCTCCTCTGCCACCTGCG
CAAGAAGCAGGGCGCCGCCAGGGCCAAGATGGAGTACAAGTGCAGCGGCCCTTCCAAGGAGGTA
GTGCTGCAGCACGTGCGGACCGAGCGGACCGCGAGAGACTCTGAGCGGCCCTCCGTCCAGGAGCC
TGGCTCCGTCCAGGAGCCTGTGCCTCCTACCCCCAGCTTTGCTACCAAAGCACCTTAGCTGGCAT
TACAGCTGGAGAAGACCCCTCCCCGCACCCCCCAAGCTGTTTTCTTCTATTCCATGGCTAACTGGCG
AGGGGGTGATTAGAGGGAGGAGAATGAGCCTCGGCCTCTTCCGTGACGTCACTGGACCACTGGGC
AATGATGGCAATTTTGTAAACGAAGACACAGTCCGATTTGTCCAGGTCTCACTACCGGGCGCA
GGAGGGTGAGCGTTATTGGTCCGCGAGCCTTCTGGGCGAGACCTTGACCTCGTGGGCTAGGGATGACT
AAAATATTTATTTTAAAGTATTTAGGTTTTGTTTGTCTTCTTACCTGTATGTCTCCAG
TATCCACTTTGCACAGCTCTCCGGTCTCTCTCTCTACAACTCCCACTTGTCATGTGACAGGTAA
ACTATCTTGGTGAATTTTTTTTCTAGCCCTCTCACATTTATGAAGCAAGCCCCACTTATTTCCCAT
TCTTCTAGTTTTCTCTCCAGGAAGTGGGCAACTCACCTGAGTCAACCTACCTGTGCCTGACCC
TACTTCTTTGCTCTTAGCTGTCTGCTCAGACAGAACCCCTACATGAAACAGAAACAAAACACTA
AAAATAAAAATGGCCATTTGCTTTTTCAACCAGATTTGCTAATTTATCTGAAATTTAGATTCCCAG
AGCAAAATAATTTTAAACAAAGGTTGAGATGTAAAAGGTATTAATTTGATGTTGCTGGACTGTCAT
AGAAATTACACCCAAAGAGGTATTTATCTTTACTTTTAAACAGTGAGCCTGAATTTTGTGCTGTTT
TGATTTGTACTGAAAAATGGTAATTGTTGCTAATCTTCTTATGCAATTTCTTTTTTGTATTATTAC
TTATTTTGTACAGTGTTGAAAATGTTTCAGAAAGGTTGCTCTAGATTGAGAGAAGAGACAAACACCTC
CCAGGAGACAGTTCAAGAAAGCTTCAAACTGCATGATTCATGCCAATTAGCAATTGACTGTCACTG
TTCTTGTCACTGGTAGACCAAAATAAAACCAGCTCTACTGGTCTTGTGGAATTGGGAGCTTGGGA
ATGGATCCTGGAGGATGCCCAATTAGGGCCTAGCCTTAATCAGGTCCTCAGAGAATTTCTACCATT
TCAGAGAGGCCTTTTGAATGTGGCCCTGAACAAGAATTGGAAGCTGCCCTGCCCATGGGAGCT
GGTTAGAAATGCAGAATCCTAGGCTCCACCCATCCAGTTCATGAGAATCTATATTTAACAAGATC
TGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATTCAGACTGCTTCCAATTTTCTGGAA
TACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAGGCCCTTATTTTCAAGAACTGAG
GAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGGTACACAGCTCTAGACACTGCCA
CACAGGGTCTGCAAGGTCTTTGGTTAGCTAAGCTAGGAATGAAATCCTGCTTCAGTGTATGGAAA
TAAATGTATCATAGAAATGTAACTTTTGTAAGACAAAGTTTTCTTCTATTTTGTAACTCAAA
ATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACACAGGCCAAATCCTTGCTTATGAC

FIG. 7

ATCACTTGTACAAAATAAACAAATAACAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAA

FIG. 7a